

Research and development of biomarkers and technologies for the clinical application of early cancer detection strategies

**FRED HUTCHINSON
CANCER RESEARCH CENTER**
A LIFE OF SCIENCE

**Early
Detection
Research
Network**

JPL

Jet Propulsion Laboratory
California Institute
of Technology

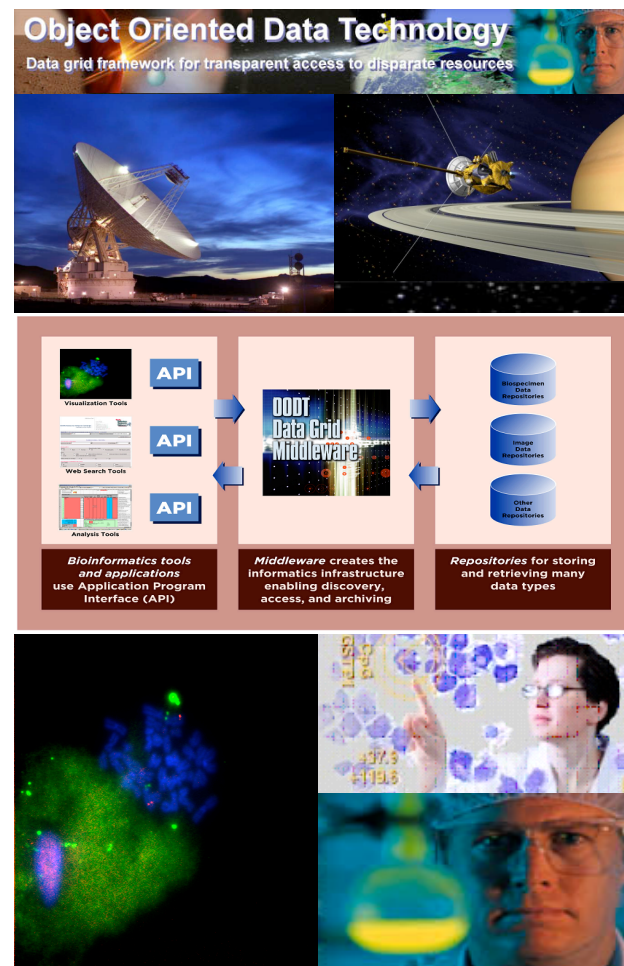
NCI EDRN Biomarkers Knowledge Environment Informatics Infrastructure

Sudhir Srivastava

*Chief, Cancer Biomarkers Research Group
Division of Cancer Prevention
National Cancer Institute*

Dan Crichton

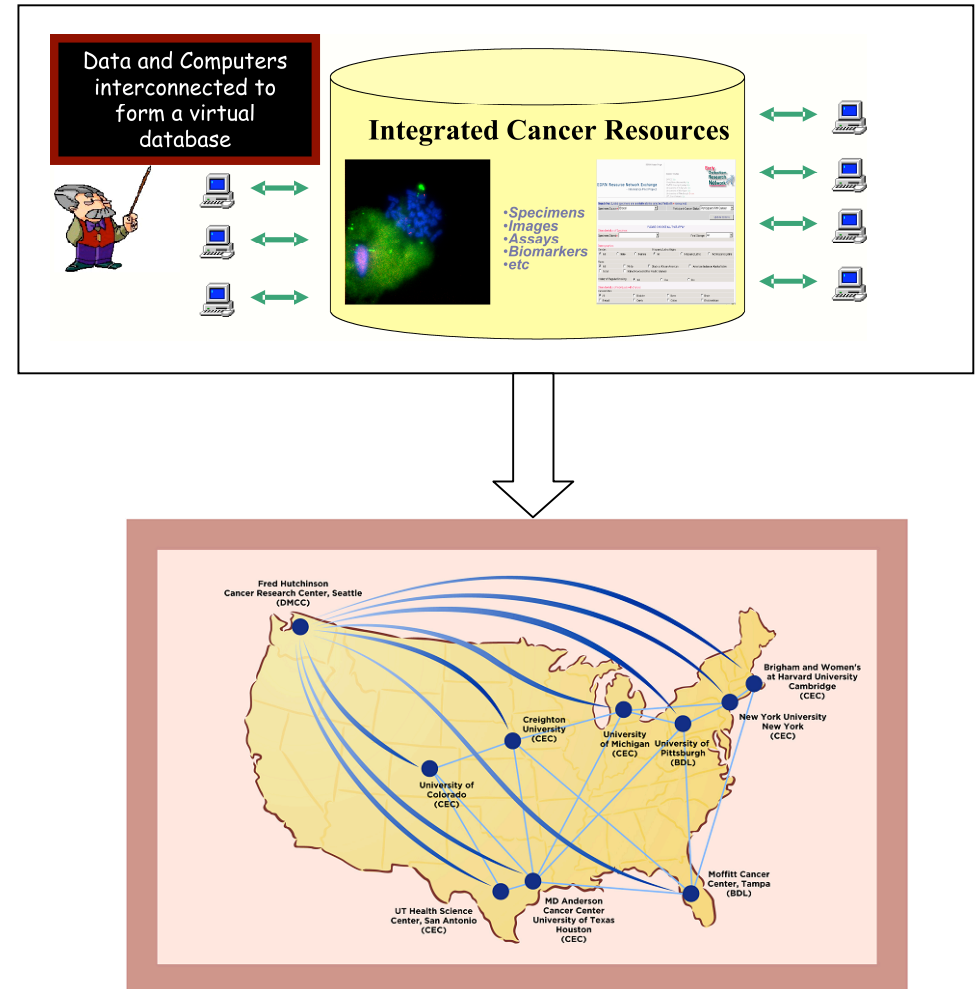
*Program Manager for Engineering, NASA Planetary Data System
PI, Informatics, NCI EDRN Program
Jet Propulsion Laboratory*



July 28, 2007

The Informatics Challenge...

- Distributed scientists across ~40 research centers
 - Parallel and on-going research
 - Diverse data sets in highly distributed catalogs
- Science discovery through capture, mining and correlation of diverse data sets acquired during EDRN validation studies
- Sharing of data resources between diverse, distributed science research databases
 - Biomarkers
 - Proteomics
 - Biospecimens
 - Various technologies and data products (image, micro-satellite, ...)



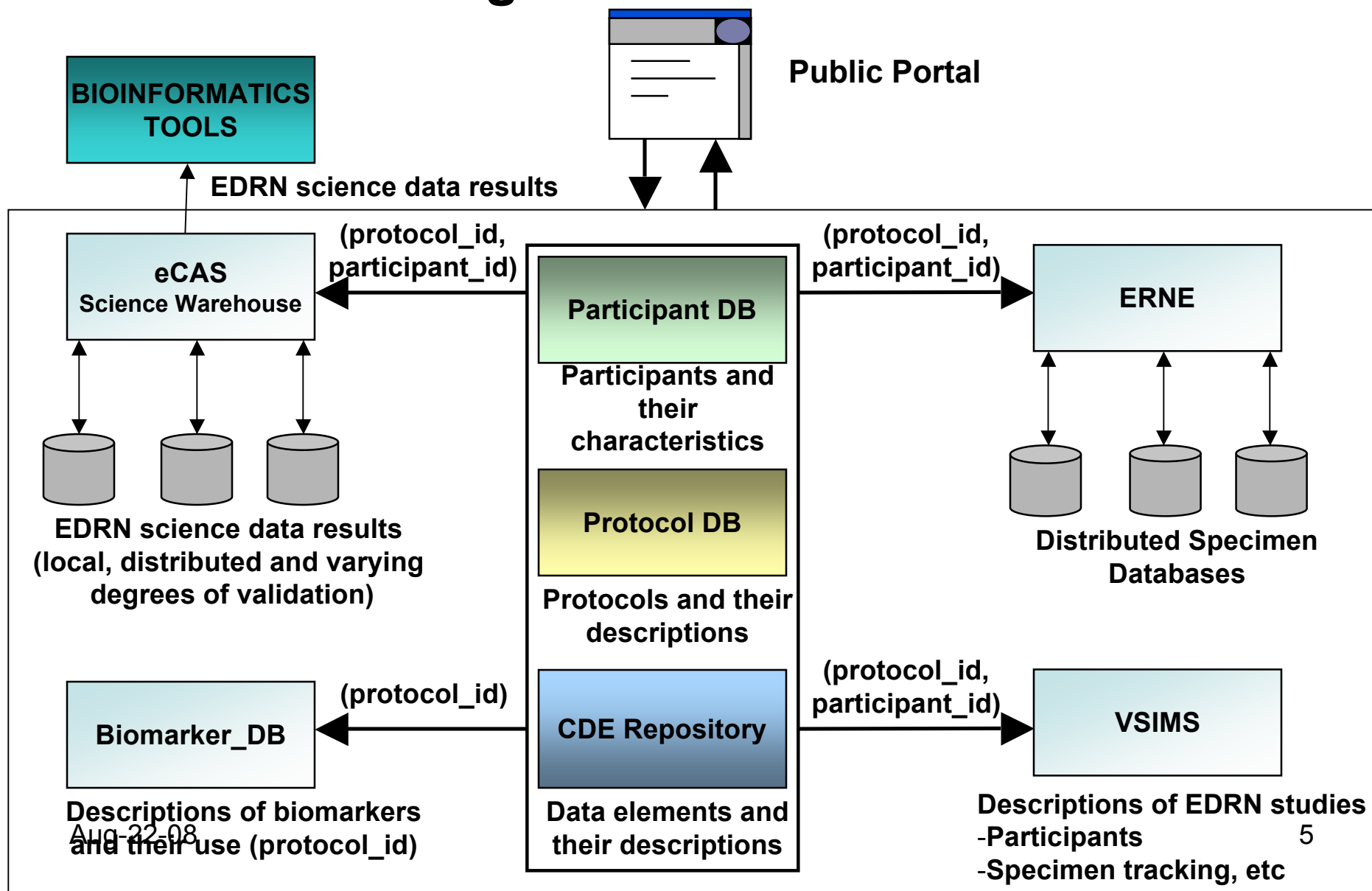
Original Informatics Vision for EDRN

- Advances in distributed computing infrastructures provide
 - Connectivity of computers and repositories globally over the Internet
 - Electronic sharing of data
 - Web-based access to data distributed in independent databases
 - Software and data standards for the access and sharing of data
 - Global security mechanisms to protect data sharing and access
- In effect, EDRN was pursuing a new paradigm for biomedical research in which data and computing can remain distributed, but be integrated into a virtual knowledge environment

EDRN Informatics Goals and Principles

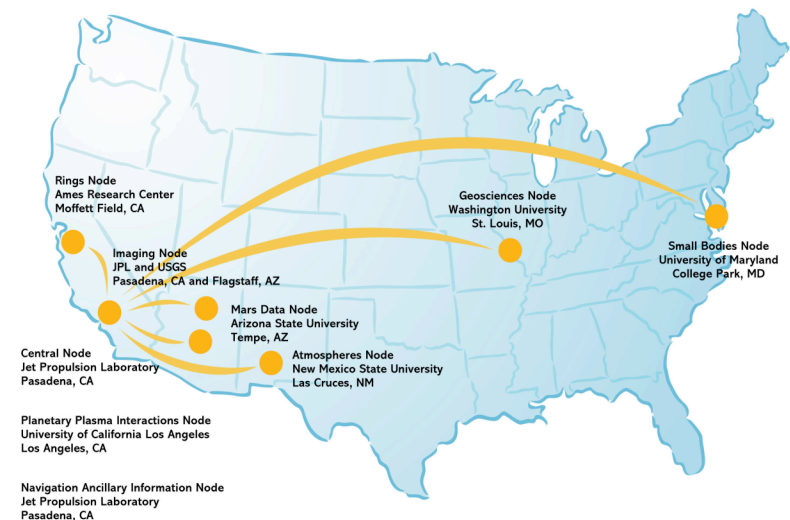
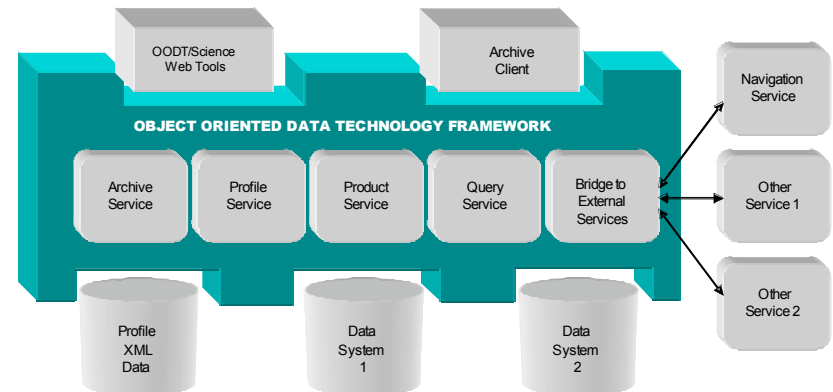
- Develop a knowledge system that **links** together EDRN data assets into a virtual data system based on **common data elements**
- Establish an EDRN bioinformatics program that promotes the use of a **common informatics infrastructure** by EDRN sites.
- Provide an infrastructure for **capturing** EDRN validation study results and a mechanism for **distribution**
- Define data and software **standards** for EDRN informatics systems
- **Collaborate** with both EDRN and non-EDRN sites on informatics.
- Develop a public portal that provides **information dissemination** about EDRN programs and progress.
- Enable tools that support **scientific inquiry** both within and across databases and data sets.

Realizing the Informatics Vision



Leveraging Software from NASA: Object Oriented Data Technology

- Started in 1998 as a research and development task funded at JPL by the Office of Space Science to address
 - Application of Information Technology to Space Science
 - Provide an infrastructure for distributed data management
 - Development of a set of software components to support generation and sharing of data for distributed science domains
- OODT Initial focus on capturing and sharing data across distributed science data repositories and systems
- Runner-up NASA Software of the Year, 2003
- Used in planetary, astrophysics earth and biomedical sciences
- Java framework available via Open source at <http://www.openchannelsoftware.com>



Sharing Specimens: A start at building the infrastructure

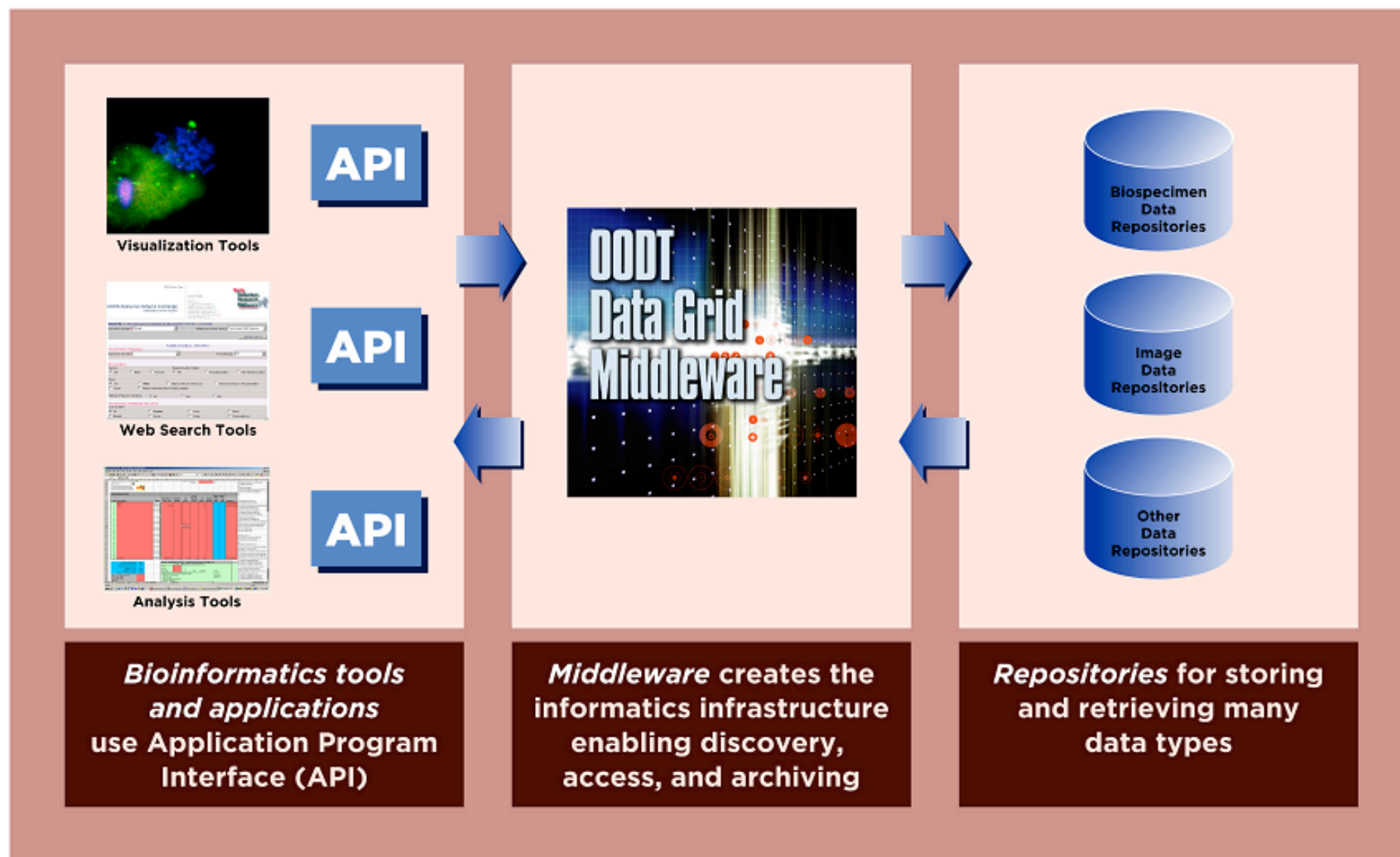
- Specimen management is a critical function that is carried out across biomedical research centers
 - Capturing the characteristics of the specimens (e.g., their epidemiological characteristics) is critical
 - Scientists need access to certain specimens during an assay or study
- Access to a “virtual specimen bank” viewed as important in supporting EDRN’s need for collaboration within the network
- But, challenges remained...
 - No standards for how specimen information is represented...often homegrown
 - Very different levels of technology, support, etc at sites
 - Specimen management is reviewed and controlled by Institutional Review Boards (IRBs) put into place by the federal government



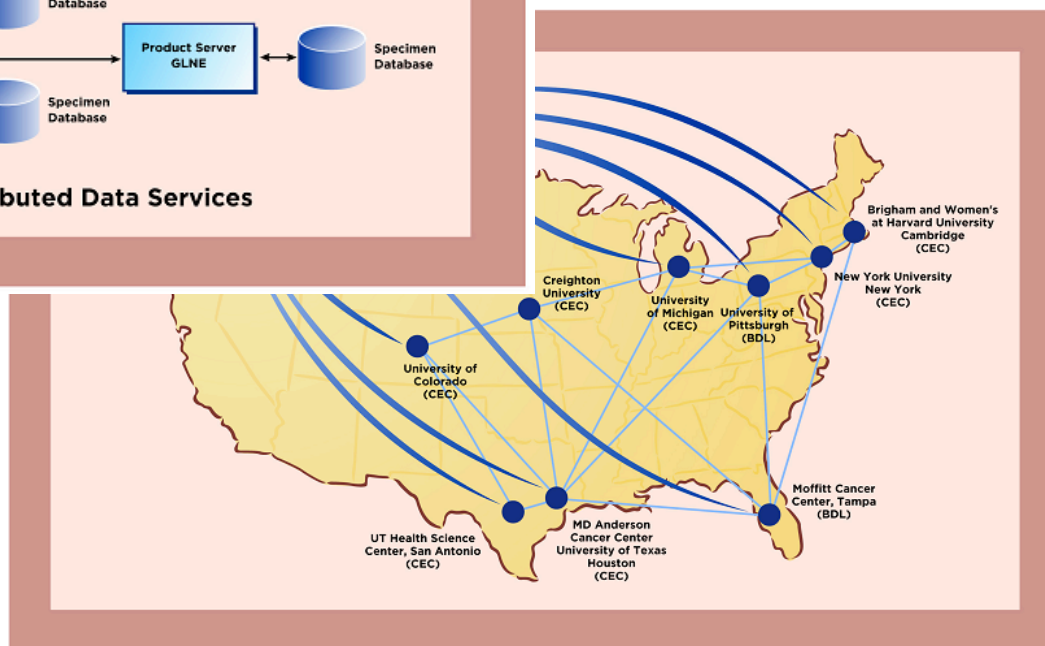
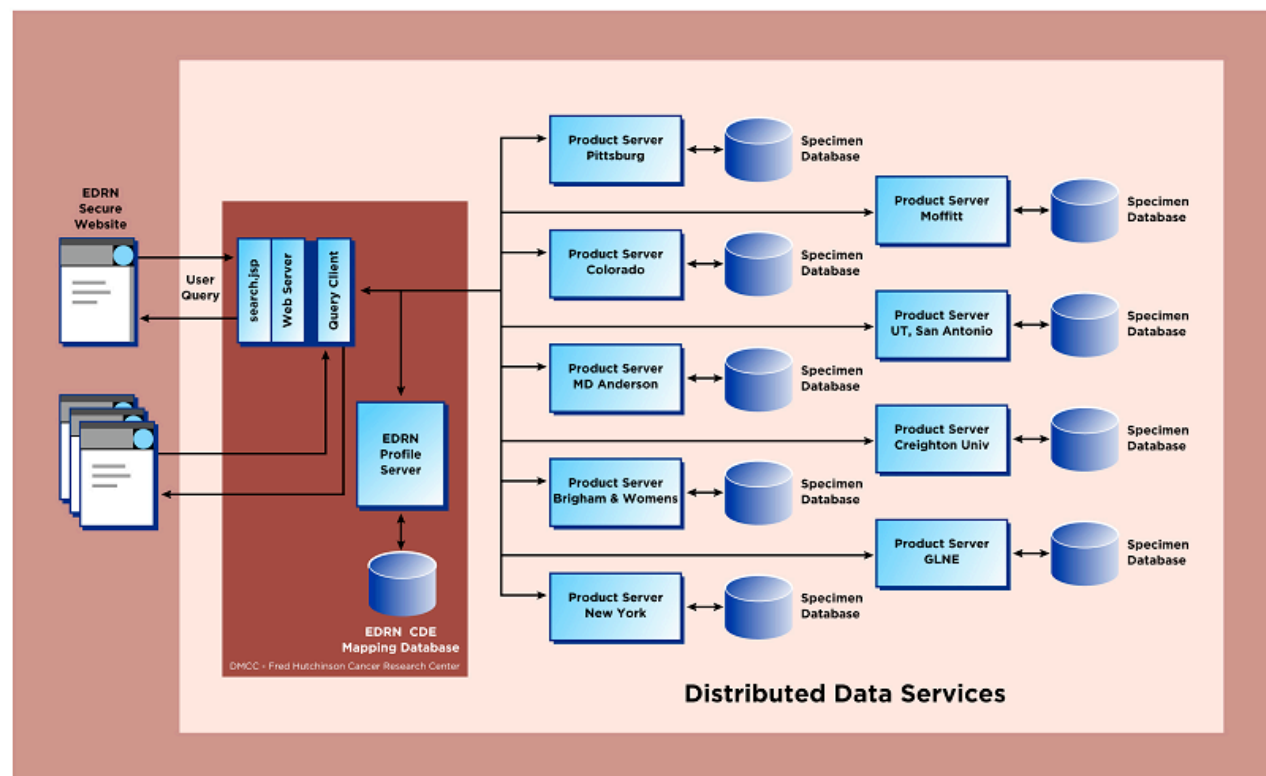
Project 1: EDRN Resource Network Exchange (ERNE)

- Initiated as a proof of concept to construct a virtual specimen system linking databases together from major cancer centers
 - Linked over 10 sites (on our way to 15)
 - Developed in three phases starting with just two sites
- Established a cross-disciplinary team including
 - Scientists (from cancer centers)
 - Policy Makers (from NCI and NIH)
 - Coordinators (from Fred Hutchinson Cancer Center in Seattle/EDRN Data Management and Coordinating Center)
 - Technologists (from JPL)
- Became an NCI Success Story...
 - 2002 Article in Journal of National Cancer Institute
 - Initiated National Biospecimen Network (NBN)

Distributed Access Model

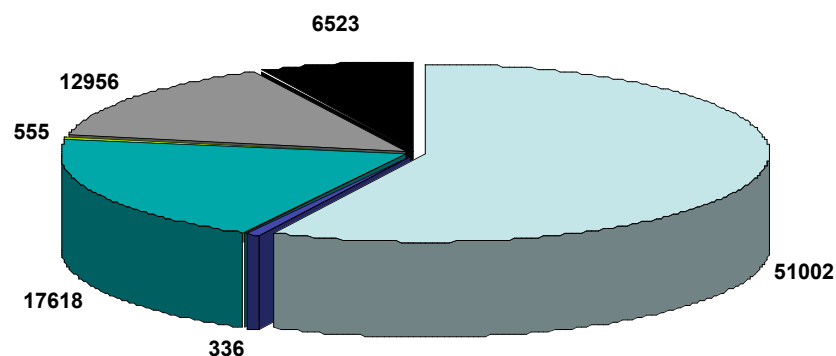


ERNE Specimen Integration and Deployment



Aug-22-08

Types of Specimens in ERNE



- 51002 - Blood
- 336 - Bone Marrow
- 17618 - Tissue
- 555 - Bronchial Washings
- 12956 - Sputum
- 6523 - Urine

Search Results - Microsoft Internet Explorer

Address: <https://ginger.hcr.org/edrn/pilot/QueryServlet?d=up&cr=up&sh=up&co=up&m=up&p=up&s=up>

EDRN Home Page

EDRN Resource Network Exchange
- Informatics Pilot Project

Options for another Search:

Search Results:

Specimen Source: Blood
Specimen Stored: DNA
Participant Cancer Status: Participants With Cancer

Protocol ID	Site ID	Site Name	# of Samples	# of Pts	Contact	Summary	Details
65	80	Creighton Univ.	288	39	patrice@creighton.edu	Summary	Details
64	72	Univ. of Colorado	3	3	wilbur.franklin@uchsc.edu	Summary	Details
36	67	Univ. of Michigan	12	3	dbrenner@umich.edu	Summary	Details

Data Summary (Creighton University, Protocol ID: 65) [Details](#)

CDE Category	Number of Samples	Number of Participants
Gender		
Male	207	14
Female	81	25
Race		
White	274	36
Unknown	14	3
Smoking History		
Smoked Regularly	49	14
Not Smoked Regularly	20	10
Unknown	219	15

Data Summary (University of Colorado, Protocol ID: 64) [Details](#)

CDE Category	Number of Samples	Number of Participants
Gender		
Male	3	3
Female	3	3
Race		
White	3	3
Smoking History		
Smoked Regularly	3	3

Data Summary (University of Michigan, Protocol ID: 36) [Details](#)

CDE Category	Number of Samples	Number of Participants
Gender		
Male	8	2
Female	4	1
Race		
White	2	1
Black or African-American	4	1
Unknown	6	1
Smoking History		
Not Smoked Regularly	8	2
Unknown	4	1

If you have any questions or suggestions, please send a feedback message to [DMCC](#).

[Members](#) | [Committees](#) | [Collaborative Groups](#) | [Resources](#) | [Informatics](#) | [Protocols](#) | [Policies](#) | [Publications](#) | [Home](#)

Download Acrobat Reader | Change Password | Registration | Privacy Policy | Upgrade Browser | Security

©Secure site maintained by COMPASS, Fred Hutchinson Cancer Research Center.
For comments and suggestions regarding this website, Email edn@photo.org

Project 2: Common Data Elements

- Common Data Elements provide a set of standard terms and values for a domain
 - They are classified into organ, epidemiological and specimen CDEs
 - Critical to getting ERNE to work**, but have historically been forms-based (lack an overarching information model)
 - Based on ISO/IEC 11179 (standard for data elements)
- Captured by EDRN and maintained by the EDRN Data Management and Coordinating Center in Seattle

Data Element Mapping Tools - Netscape

URL: <https://www.compass.fhcr.org/DEMMapping/mapping/bin/DESearch.asp>

Early Detection Research Network

SEARCH FEEDBACK SITE MAP EDRN PUBLIC

Relevant Sites: [Dropdown]

Staff Name: Mark Thornquist
Site Name: Fred Hutchinson Cancer Research Center (DMCC)

Search Criteria:

- DE Definition: Race
- Search Method: Matches on all words (AND)
- DE Context: EDRN
- DE Object Class: All
- DE Property: All
- DE Phase: All
- DE Version: All
- DE Form: All
- Permissible Value: All
- Current Version: Yes (Current Version Only)

DE Name: PPT_DEMOGRAPHICS_RACE_CODE

Document Text: Race (What is your race? Check all that apply.)

Context:

Object Class: Participant

Property: Demographics

Version: 1.0

Data Type: Character

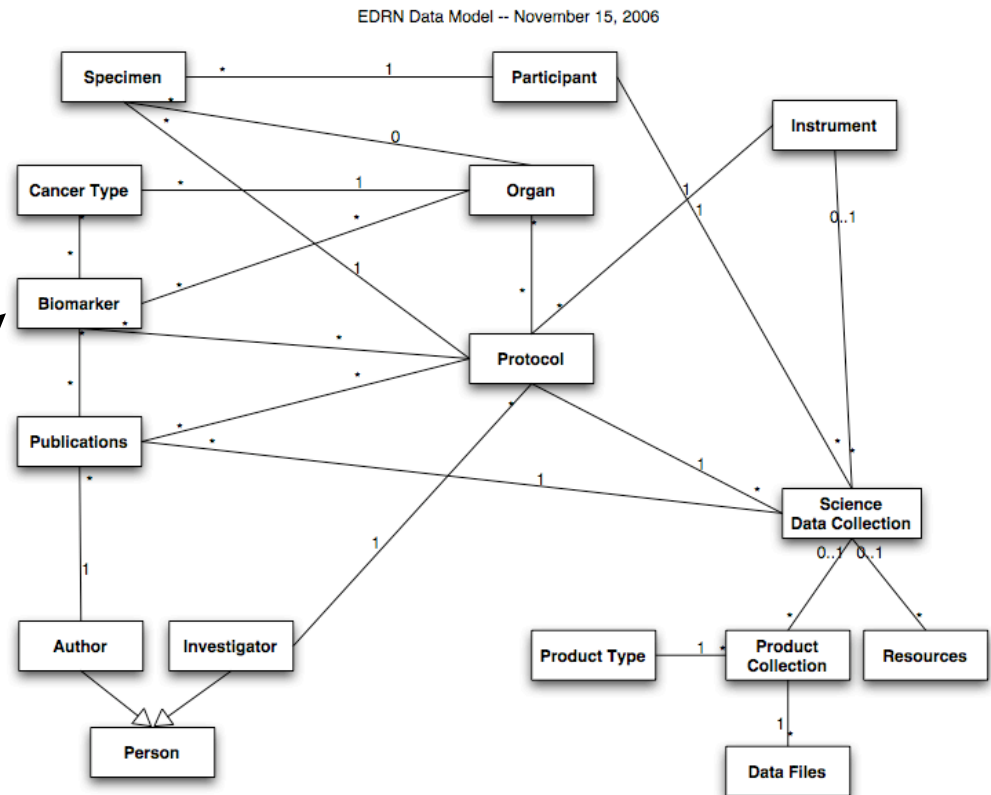
Permissible Value: List ID	List Value	Source
1	White	NCI
2	Black or African-American	Metathesaurus
3	American Indian or Alaska Native	NCI
4	Asian	Metathesaurus
7	Native Hawaiian or other Pacific Islander	NCI
95	Other	Metathesaurus
97	Other, specify:	NCI
99	Unknown/refused	Metathesaurus

Note: Permissible Values listed above are specified for all related forms.

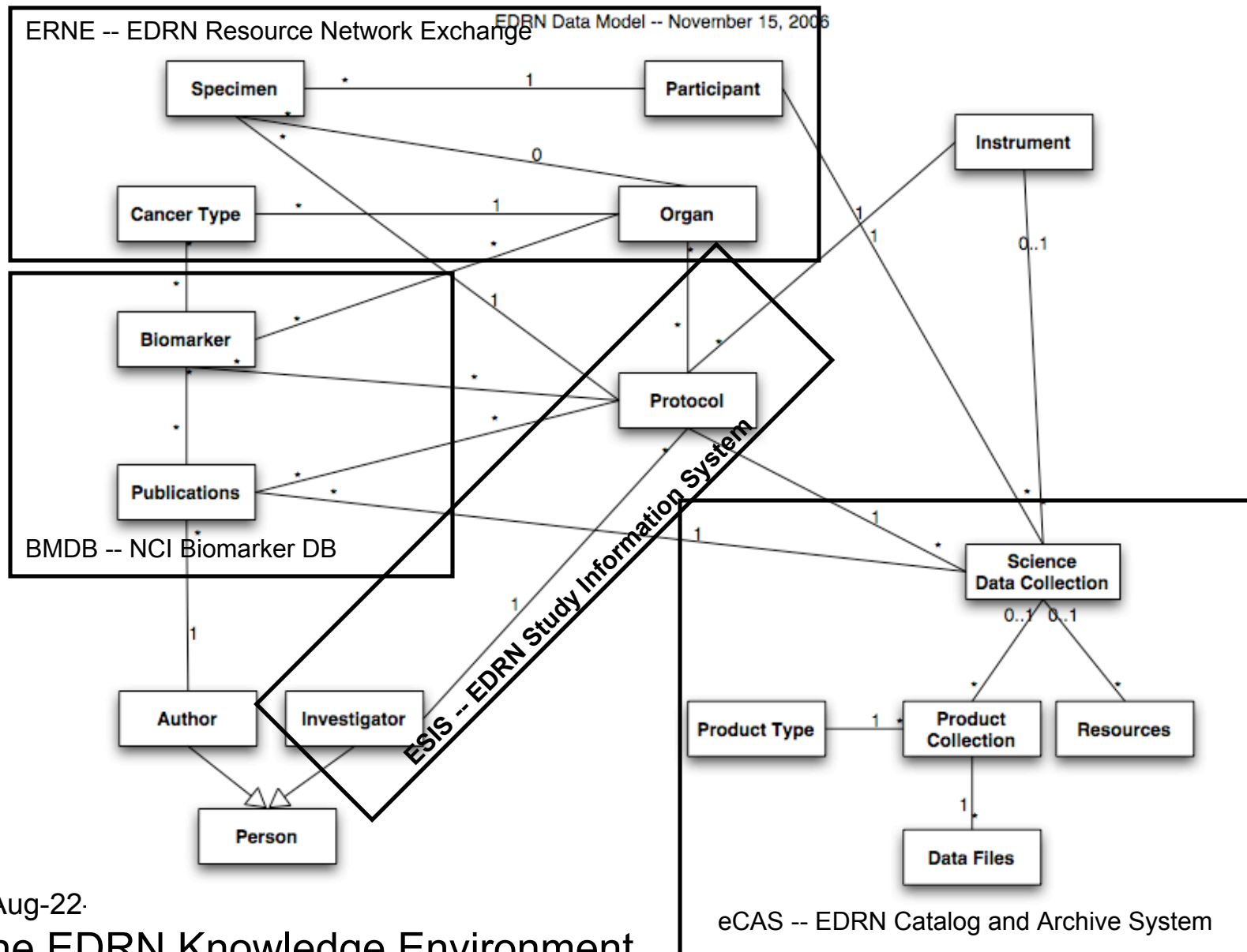
Forms Applied: [Breast Reference Set Participant Questionnaire \(1.0 - Approved\)](#)
[Prostate Reference Set \(1.0 - Feedback\)](#)
[Required Baseline Luna Ref Set \(1.3 - Initialization\)](#)

Project 3: EDRN Information Model

- High level *ontology* model of EDRN which describes
 - Core “object” concepts of EDRN data
 - Relationships between those objects
- Specific models are derived from this high level model
 - Model of biospecimens
 - Model for each class of science data
- EDRN is specifically focusing on a granular model for annotating biomarkers and their studies
- But, the high level model allows us to move towards an integrated information environment
- **The model is independent of the software component architecture**



EDRN Data Model Mapping to Applications

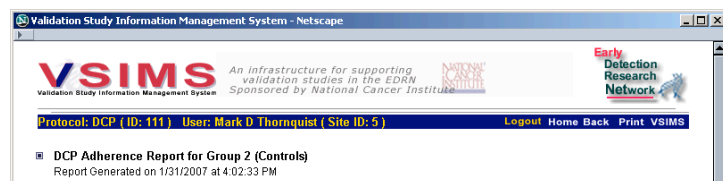


Aug-22:
The EDRN Knowledge Environment

Project 4: Validation Study Management

- **VSIMS:** Online study management system supporting all EDRN validation studies. Built upon the EDRN CDE repository and using reusable modules to speed development for new studies.

eSIS: System in development to track the progress of all EDRN-funded projects, including timelines, GANTT charts, phases of development, current study status



VSIMS instance
for DCP validation
study

Site	Total	Number Due	D S V
143 - University of Michigan	80	65	11
144 - Mount Sinai Hospital	72	56	3
145 - Mayo Clinic	27	24	2
146 - Saint Louis University	40	28	5
147 - Stanford University	60	50	1
148 - University of Pennsylvania	49	35	3
275 - Mayo Clinic Jacksonville	12	6	6
Total for All Sites	340	264	4

Total represents the cumulative number of participants confirmed. This equals the sum of 'Total Number Completed In Window' and 'Completed Out of Window'. **Completed In Window** represents those participants who completed out of window - either too early or too late. **Not Done** represents those participants who have not nor

Note: Click the link to view the Detail List.

Home Back

Secure site maintained by COMPASS, Fred Hutchinson Cancer Res

VSIMS An infrastructure for supporting validation studies in the EDRN Sponsored by National Cancer Institute

Protocol: DCP (ID: 111) User: Mark D Thornquist (Site ID: 5) Logout Home Back Print VSIMS

Submit Data
Confirm Eligibility
Specimens
Reports
Issue Tracking
Data Transfer
Study Info
Administration

Study Update - DCP:
The EDRN Validation Study entitled "Validation of Serum Markers for the Early Detection of Hepatocellular Carcinoma" will investigate a new des gamma carboxyprothrombin (DCP) for liver

Documents:
DCP Newsletter - N
DCP Newsletter - S
DCP Newsletter - J

Recruitment Plots
New recruitment/month
Cumulative recruitment

Recruit a Matched Control:
Click here to determine if a potential Control has a matching Case.

Message Board
New DCP Protocol and Manual of Operations!
Please read Version 1.7 of the Protocol and version 1.4 of the Manual of Operations released October 12, 2006.
The revision histories explaining the changes are documented in Bulletins #17, 18, and 19.

Secure site maintained by COMPASS, Fred Hutchinson Cancer Research Center®. Last updated on 1/31/2007. Contact: edmdmcc@thore.org

EDRN PROTOCOL
Search, Register, Update and Report EDRN Protocol

Search Back to: Web Site List

Please select a protocol name to get its information in detail.

Protocol Name	Protocol ID
Androgenic Activity and the Association with Ovarian Cancer	Unregistered
Assessment of Clinical Potential of Digital SNP Analysis	Unregistered
Barnett's Esophageal Methylation Profiles	Unregistered
Biomarker Dev Lab A1 Moffitt Supplement	72
Biomarker Development Laboratory	95
Biomarkers in Familial Multiple Myeloma	139
Body Fluids as a Source of Diagnostic Biomarkers	76
Body Fluids as a Source of Diagnostic Biomarkers: Breast	68
Body Fluids as a Source of Diagnostic Biomarkers: Prostate	67
Breast Ductal Lavage Fluid Specimen Bank	Unregistered
Breast tissue array	71
Clinical Validation of Molecular Signatures of Cervical Neoplasia	60
CLUE Studies: Evaluating Biomarkers of Carcinogenesis	59
Collaborative study on SELDI biomarker discovery and serum protein profiling of lung cancers	Unregistered
DCP Validation Study	Unregistered
Detecting breast cancer protein signatures in body fluid	Unregistered
Detecting Breast Cancer Signatures in Body Fluids	73
Detection of Bladder CA by Microsatellite Analysis (MSA)	108
Detection of cancer DNA markers in urine	Unregistered
Detection of Serum Markers of Breast Cancer by SELDI	Unregistered
Detection of Serum Proteomic Patterns and SNPs in Breast Cancer: A Pilot Study	112
Development and Evaluation of a Tumor Marker for Prostate Cancer	57
Development of FISH-based assay for early detection of high grade cervical dysplasia	96
Development of paraffin-based expression profiles to predict minimal vs. substantial risk of disease progression using retrospective datasets	113
Early detection of liver cancer and hepatitis	97
Early Detection of Urinary Bladder Cancer	Unregistered
Early Detection Research Network Registry for Hereditary Cancer	28
Early Molecular Detection of Invasive Bladder Cancer by Urine Examination	Unregistered
Ecogenetic Study of Lung Cancer	105
EDRN Informational Data Element	CR

Aug-22-08

15

Project 5: Biomarker Database

- Develop a registry to annotate biomarkers that are either under development or reported in publications
 - Entry into the registry would be through peer review
 - Initiated by EDNRN, but open to NCI

Biomarker Descriptions

SELDI Protein Profiles — EDNRN Demonstration Site

http://edrndemo.org/bmdb/biomarkers/seldi-protein-profiles

Biomarker Summary Detail

SELDI Protein Profiles (see related studies, science data, or specimens) → BioDB home Related Resources Prostate Cancer Foundation

Protein profiles from two independent laboratories using SELDI-TOF-MS and IMAC Proteinchip, using serum from 500 prostate cancer cases (250 with aggressive/advanced disease and 250 with intermediate- or low-risk disease) as well as 250 prostate cancer controls (biopsy-negative with a range of PSA values). An additional group of 50 patients with other cancers but no evidence of prostate cancer and 50 patients with various inflammatory diseases will also be examined.

Prostate Urethra

Phase	2	
Classification	Protein	
QA Status	Not ascertained	
Sensitivity Range	36.9 → 46.9	No comment
Specificity Range	46.9 → 46.9	Effective
Positive Predictive Value Range	12.9 → 23.4	Expected range
Negative Predictive Value Range	-12.9 → -9.2	Narrow range
Assays Used	serum protein profile	
Technologies Used	SELDI-TOF-MS, IMAC Proteinchip	

Publications

Title	Pub Med ID	Affiliation	Authors
Evaluation of Serum Protein Profiling by Surface-Enhanced Laser Desorption/Ionization Time-of-Flight Mass Spectrometry for the Detection of Prostate Cancer: I. Assessment of Platform Reproducibility	15613711	University of Alabama, National Cancer Institute, Virginia Prostate Center, Hillman Cancer Center	William Grizzle, Dean Brenner

Biomarker Studies

SELDI Validation Study — EDNRN Demonstration Site

http://edrndemo.org/bmdb/protocols/seldi-validation-study/

Early Detection Research Network

Home About EDNRN EDNRN Investigators Committees Organ Groups Resources Publications

Home → Biomarker Database → Protocols → SELDI Validation Study

Biomarker Study Detail

SELDI Validation Study (see related science data or specimens) → BioDB home Related Resources Prostate Cancer Foundation

A Comprehensive Program for the Validation of Prostate Cancer Early Detection with Novel Protein Identification Techniques -- is divided into three phases. The goal of Phase I was to assess the reproducibility and portability of Surface-Enhanced Laser Desorption and Ionization time-of-flight mass spectrometry (SELDI-TOF-MS) using protein profiles generated from serum. Phase I was recently successfully completed at six institutions using a single source of pooled sera. The overall goal of Phase II is to develop and evaluate an algorithm for classifying cases and controls using protein profiles produced from SELDI-TOF-MS using serum collected from prostate cancer cases and non-cancer controls.

Prostate Urethra

From biomarker research SELDI Protein Profiles.

Phase	2	
Sensitivity Range	36.9 → 46.9	No comment
Specificity Range	46.9 → 46.9	Effective
Positive Predictive Value Range	12.9 → 23.4	Expected range
Negative Predictive Value Range	-12.9 → -9.2	Narrow range
Assays Used	serum protein profile	
Technologies Used	SELDI-TOF-MS, IMAC Proteinchip	

- Aug-22



Current Status of Biomarker DB

- V1.0 of the model is complete
 - Capture of information is now underway
- A pilot database has been developed and integrated into the EDRN science portal (more on that later...)
 - Will provide a beta test release in September as part of the EDRN Public Portal
- EDRN is establishing assembling a curation process and group for the database

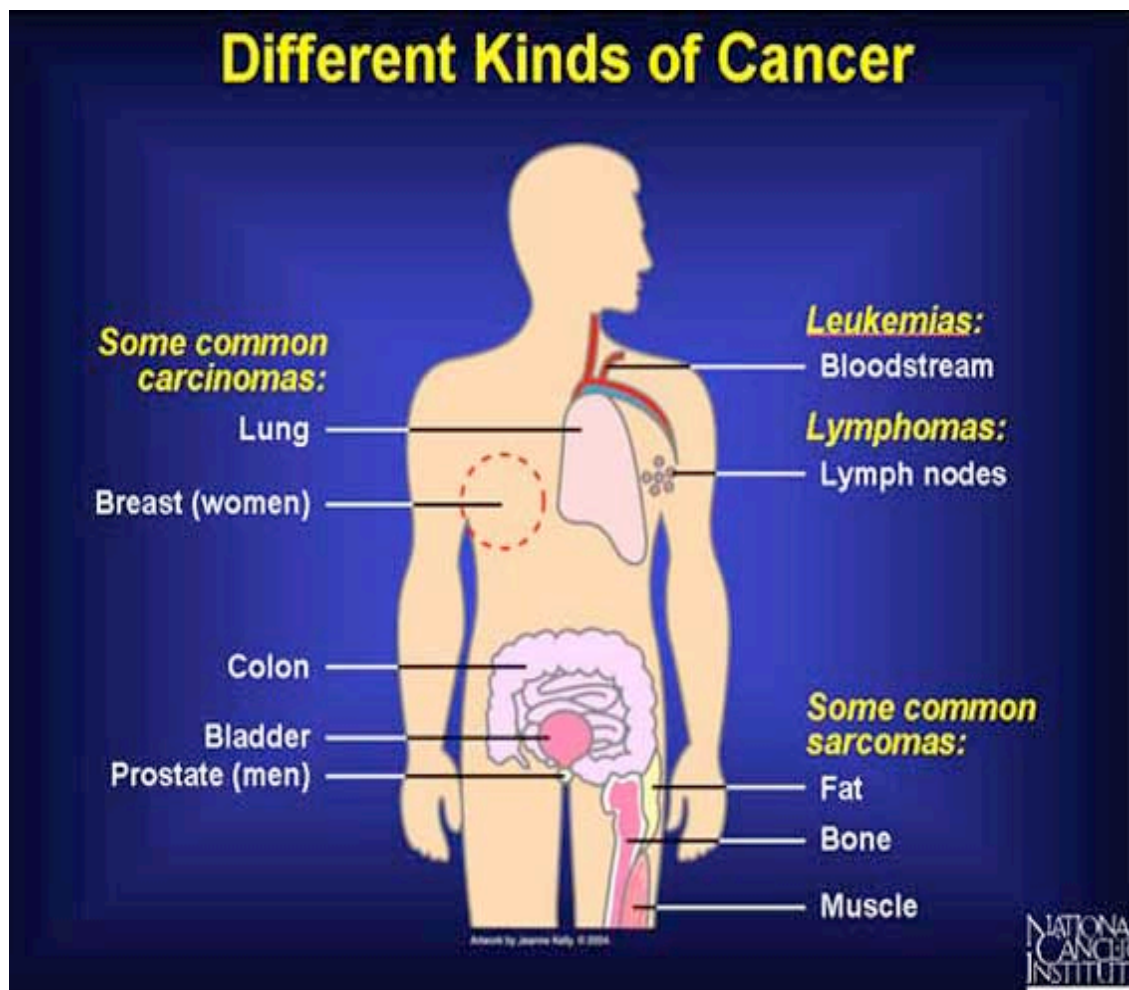
Project 6: Science Data Warehousing and Access

- Provide an integrated repository of EDRN data sets (e.g., mass spectrometry data, microarray data, 2-D electrophoresis gels, etc.) released by investigators
 - Use the EDRN CDEs to populate a catalog describing the data sets
- Provide tool set for constructing (managing and generating) biorepositories
- Provide a distribution mechanism to the community for EDRN public science data
 - Provide granular searches across distributed data sets, integrated into the EDRN knowledge system
- Provide long term preservation of EDRN study information

EDRN Science Data Covers Different Types of Cancer

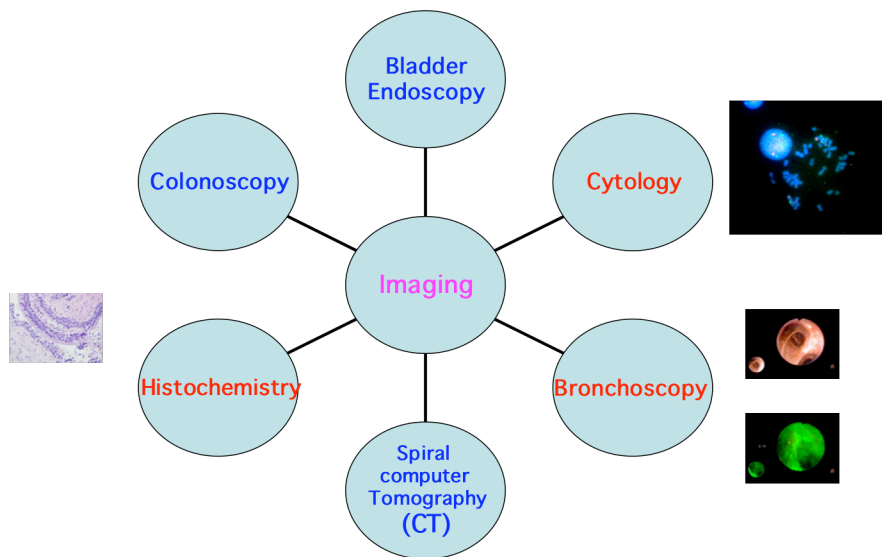
Common Cancer Types

Bladder Cancer
Breast Cancer
Colon Cancer
Endometrial Cancer
Kidney (Renal Cell) Cancer
Leukemia
Lung Cancer
Melanoma
Non-Hodgkin's Lymphoma
Pancreatic Cancer
Prostate Cancer
Skin Cancer (Non-melanoma)
Thyroid Cancer
Ovarian Cancer
Liver Cancer



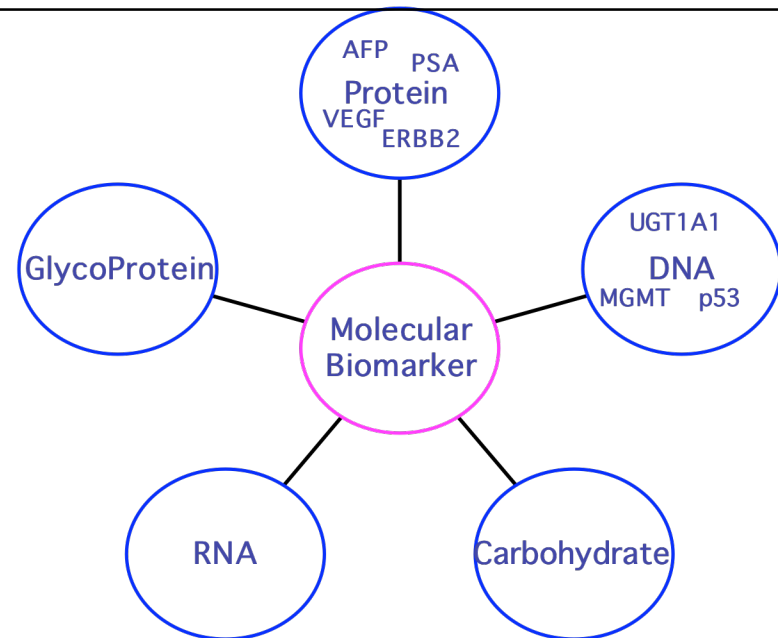
Managing EDRN's Multi-disciplinary Science Data

EDRN's underlying information model needs to be flexible in order to manage a variety of different types of science data captured by different experiments at different phases in studying biomarkers...



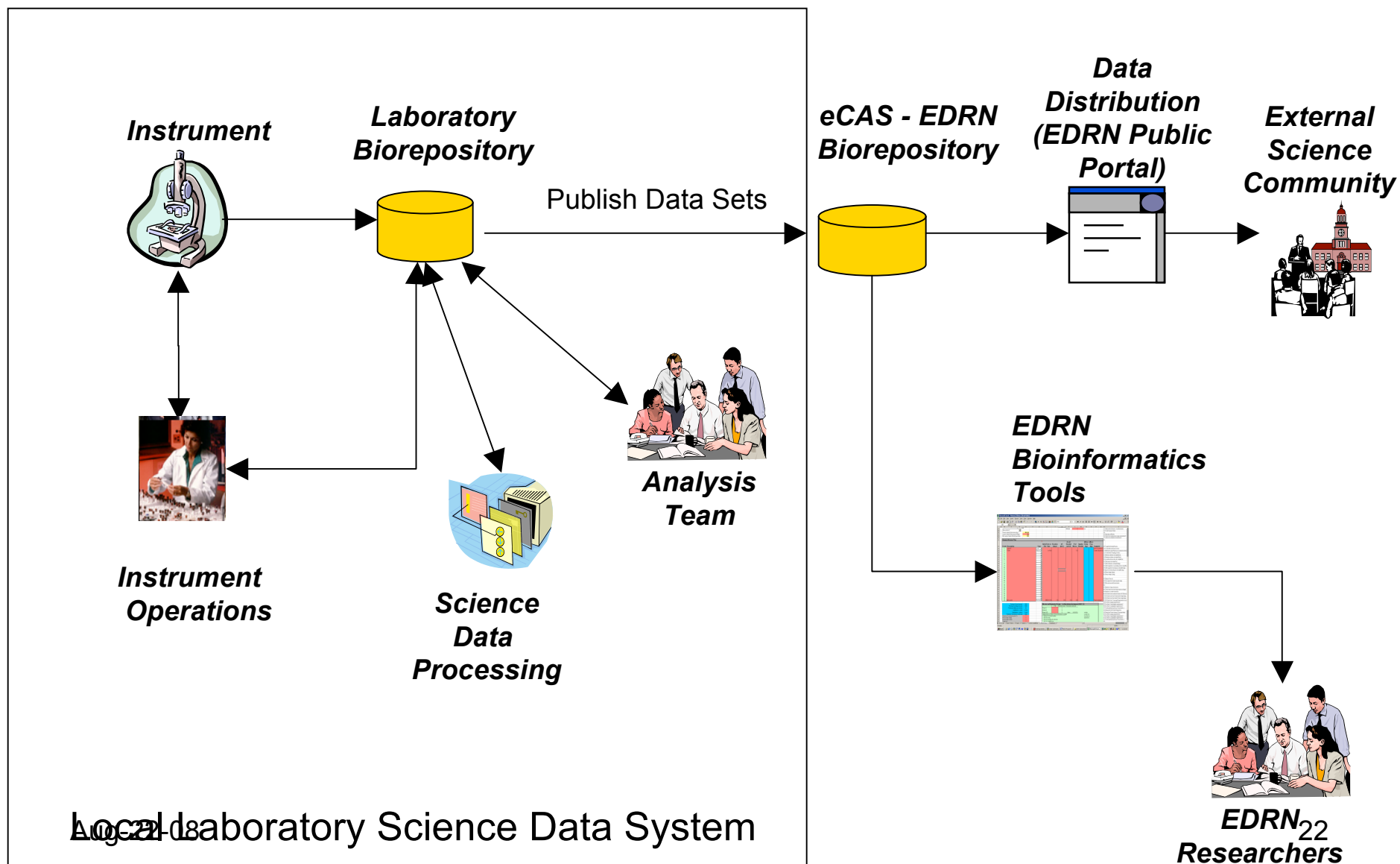
Aug-22-08

Imaging Data Types



Different Data Types

EDRN Science Pipeline



The EDNRN “Biomarker Atlas”

- Focus on organ centric discovery and access to science data
- Distributed Biomarker Atlas for the Lung
 - Prototype with access to distributed lung image databases (Colorado and Roswell Park)
 - Registration of science products against lung map
 - Integration of distributed image and specimen data sets
 - Demonstrated at the Lung SPORE meeting in July by Wilbur Franklin

Aug-22-08

Early Detection Research Network

Home About EDNRN EDNRN Investigators Committees Organ Groups Resources Publications

Home

Biomarker Search: Lung

expand (+)

Patient Characteristics

Gender: ☒ Male ☐ Female

Ethnicity:

☐ Caucasian

☐ African American

☐ American Indian or Alaska Native

☐ Native Hawaiian or other Pacific Islander

☐ Asian

☐ Other

☐ Unknown / refused

Born between 1928 and 2007

Patient History of Cancer: ☐ No ☒ Yes

Type: Lung

Living in: 90210-0001 (zip code)

☐ Smoker ☒ Non-Smoker

Submit

Link: [Return to Map View](#)

Search Results

40529 R87-3F.jpg
Pathology Score: 940
Participant Id: 6227965

8171 LUL-19.jpg
Pathology Score: 832
Participant Id: 6186616

21432 L86-1.jpg
Pathology Score: 891
Participant Id: 6155419

15453 R87-3F18.jpg
Pathology Score: 833
Participant Id: 6341689

9271 RUL-3F.jpg
Pathology Score: 821
Participant Id: 6819142

8171 R88-9 - 38.jpg
Pathology Score: 310
Participant Id: 6097731

L86-1F 40243.jpg
Pathology Score: 410
Participant Id: 6096040

40222 LU08-1.JPG
Pathology Score: 891
Participant Id: 6058037

Identifier: bf8d4b46-1fb2-11dc-aa9a-030dc0903a6
Title: 21432 L86-1.jpg Specimen Collected Date: 06202007
Study Participant Id: 6155419 Bronchial Site Code: 8 Pathology Score: 891

Biomarker Atlas for Lung

Project 7: Public Portal

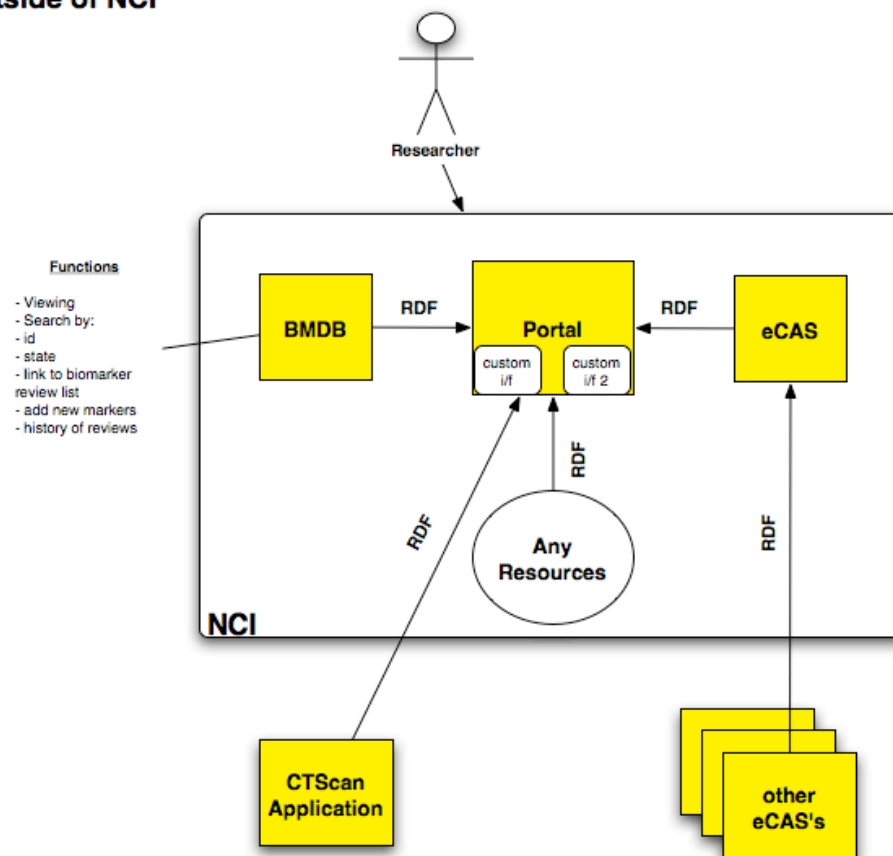
- Originally released in 2005
 - Built on Plone - open source portal software
- Used by NCI and EDRN for disseminating program information
- But, will migrate to a science portal soon...
 - Currently under development



Moving to an Integrated Science Architecture

- Semantic science portal driven by the EDRN ontology
 - Schema loaded into the ontology via RDFS (and Protégé)
 - Metadata from distributed applications dumped into the portal via RDF

Outside of NCI



Integrated EDRN Knowledge Environment



- Model-driven Portal
 - Adapts to a dynamic set of “object types” (mass spec, immunohistochemistry, etc) based on the model
 - Provides access to distributed repositories of information
 - Provides “google-like” search of the object types

The screenshot shows the EDRN Knowledge Environment website. At the top, there's a red banner with the National Cancer Institute logo and the text 'Early Detection Research Network'. Below the banner is a navigation bar with links: Home, About EDRN, EDRN Investigators, Committees, Organ Groups, Resources, Publications, and Data. A search bar is located on the right side of the navigation bar.

Below the navigation bar, the main content area is divided into several sections:

- Left Sidebar:** Contains links to various resources: Calendar, Division of Cancer Prevention (DCP), Cancer Biomarkers Research Group (CBRG), Find EDRN Protocols, Funding Opportunities, EDRN Bookshelf, and a 'Need Help?' section with contact information (1-800-4-CANCER).
- Main Content Area:**
 - Welcome Message:** A paragraph welcoming users to the EDRN Knowledge System and providing contact information for the EDRN Informatics Team.
 - Search Terms:** A section explaining how to use the search bar, including a search box and a 'Search' button.
 - Search Examples:** A list of search terms and their corresponding results:
 - metastatic marker:** Finds resources that have "metastatic" or "marker", with preference to resources with both.
 - "metastatic marker":** Finds resources with the exact phrase "metastatic marker".
 - metastatic and marker:** Use **and** to find items with both "metastatic" and "marker" in the resource.
 - metastatic or marker:** Use **or** to find items with either "metastatic" or "marker" in the resource.
 - meta*:** Finds resources that have "meta" as a prefix, such as "metastatic", "metastasis", "metastasis", etc.
- Right Sidebar:** A 'Resources' section with links to Specimens, Protocols, Biomarkers, Organs, Data, Diseases, Sites, and Publications.

Progress in cancer research towards science-driven informatics architectures

- Recognition of how to architect science-driven distributed software systems*
 - Separate the architecture into core pieces (process, data and software)
 - The “information model” is critical
 - Should provide a generalized mechanism to describe and organize data
 - Model-driven systems provide the agility to support multi-project, multi-center studies
 - Develop modular software components that can be configured based on the “information model”
 - Modularity helps to drive both longevity and agility in system designs
 - Allow for geographically distributed software components to communicate based on standards
 - Identify and implement core scientific “use cases” that help to evolve the system
 - EDRN has demonstrated this architecture can work in managing and sharing specimen information
 - JPL has done this for planetary science and is now working with international space agencies to provide access to scientific data results returned from international missions

*D. Crichton, S. Kelly, C. Mattmann, Q. Xiao, J. S. Hughes, J. Oh, M. Thornquist, D. Johnsey, S. Srivastava, L. Esserman, W. Bigbee. **A Distributed Information Services Architecture to Support Biomarker Discovery in Early Detection of Cancer.** In Proceedings of the *2nd IEEE International Conference on e-Science and Grid Computing*, pp. 44, Amsterdam, the Netherlands, December 4th- 6th, 2006.

More Lessons Learned...

- Technology needs to be accessible (with a different levels of entry)
 - We need to work with sites to help them understand technology and use it compliant with federal government regulations
 - We can't do technology for technology-sake. We need to “enable” science and demonstrate its value.
- NASA and NIH science challenges are similar
 - The science is always evolving
 - Distributed, PI-driven, complex data structures, computationally-intensive, etc
 - Need for software that supports automated pipelines
 - The need to capture and share data within scientific communities
- ***Success has been the result of having a “focused approach” to informatics and building cross-disciplinary teams...***

EDRN Informatics can pull it together...

“The most successful and efficient research about molecular markers will require effective interdisciplinary communication and collaboration involving fields of molecular biology, observational epidemiology and biostatistics.”



Ransohoff, Nature Rev Cancer 2004; 4:309-314

Aug-22-08

Acknowledgements

- Mark Thornquist and members of Data Management and Coordinator Center at the Fred Hutchinson Cancer Research Center
- Sudhir Srivastava and Don Johnsey, National Cancer Institute
- The EDRN informatics advisory group Bill Bigbee, Laura Essermann, Wilbur Franklin, Tony Hollingsworth, Jeffrey Marks
- Currently integrated sites:
 - H. Lee Moffitt Cancer Center
 - University of Texas, San Antonio
 - Creighton University
 - University of Colorado
 - University of Pittsburgh
 - University of Michigan/Dartmouth University (Great Lakes New England Consortium)
 - Brigham and Womens
 - MD Anderson
 - New York University
- NASA Jet Propulsion Laboratory